Real-Time Yoga Pose Detection Using OpenCV and MediaPipe

Isha Chaudhary
Department of Computer Science Engineering
Chandigarh Univerity
Punjab, India
ishac308@gmail.com

Nongmeikapam Thoiba Singh
Department of Computer Science Engineering
Chandigarh Univerity
Punjab, India
nthoiba12@gmail.com

Mahak Chaudhary
Department of Computer Science Engineering
Chandigarh Univerity
Punjab, India
mahic2116@gmail.com

Komal Yadav
Department of Computer Science Engineering
Chandigarh Univerity
Punjab, India
kysunita08122001@gmail.com

***Abstract*—In this era of advanced technology, where everything is going online, from shopping to education, there may be ways to get an automated, personalised trainer and mentor for yoga and meditation. Yoga is one of the best exercises for maintaining good physical and mental health. But what if yoga is done incorrectly? It can be harmful to our bodies, so there is a need for a personalised yoga trainer and mentor who is available 24 x 7 to assist you with the poses and provide you with immediate feedback in order to maintain a healthy and wealthy body. In this model, we are coming up with the idea of providing real-time posture detection using deep learning and computer vision that will guide the user to the correct postures after comparing them with the standard yoga postures and also help them track their gym activities. Our research also aims to integrate all of the features from various platforms related to posture detection applications on a single platform rather than switching between multiple platforms. In order to make it more user-friendly, we are giving users the option of selecting between an AI tutor and a physical trainer based on their needs. One of the objectives of the model is to display information about the pose, such as the benefits of performing it and how to incorporate it into routine exercises that the trainee can simultaneously practise to increase its effectiveness.**

Keywords— yoga, MediaPipe, OpenCV, pose detection

I. INTRODUCTION

Yoga is an ancient practise originating in India that encompasses physical, mental, and spiritual elements. It involves various postures that offer different benefits. With the rise of technology and reliance on virtual communication, unhealthy habits such as poor eating and sleeping patterns have become more common, posing a threat to individuals' physical and mental health. Therefore, incorporating yoga into our lifestyle can play a crucial role in promoting productivity. Identification of human activity involves creating a model of the human body that can distinguish an individual's distinct movements and actions using data obtained from sensors or vision-based methods [1]. The model must be able to recognise basic movements like sitting, standing, talking, and walking as well as more intricate and complex gestures such as various yoga poses or specialised athletic positions. Yoga assists us in understanding our bodies and achieving a state of balance between our minds and bodies. It stimulates internal organs and glands, lowering disease risk. One of the most significant benefits of yoga is its ability to relieve stress, fatigue, and invigorate the body while also boosting immunity and promoting peace of mind [2]. Consistently practising yoga has no side effects, but it is essential to perform the postures correctly to avoid injuries. To ensure proper form, we may need a trainer to guide us. However, finding a physical trainer can be challenging with our busy schedules. Therefore, we propose an AI-powered machine learning product that can recognise yoga postures and provide customised feedback to help users improve their form. Deep learning and computer vision are scientific fields that aim to model human behaviour and interpret actions. Human activity recognition involves modelling unique human motions and behaviours using data from sensors or vision-based techniques. This paper focuses on accurately identifying yoga postures, even if they are similar, by utilising features such as angles between body parts. Fig. 1 depicts the basic structure of the system.

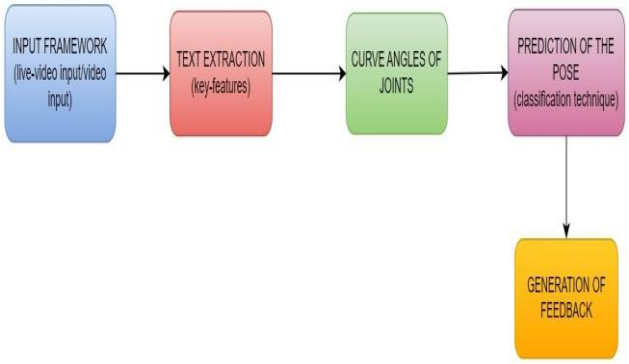

Fig. 1. Basic structure of system

Classification networks such as multilayer perceptrons and hyperparameter tuning are used to improve accuracy. There is a need for a real-time trainer who is always available to the user and who can assist in detecting yoga poses, as practising the correct yoga poses is very important as doing yoga with incorrect postures can cause various inverse effects that can be injurious to health, both physically and mentally.

II. RELATED WORK

As part of the literature review, we looked at several applications of the yoga pose detection system and the related

work that has been done in the market. In addition, we looked into different technologies that had been used in previous work to detect yoga postures. Human activity detection has been used in a wide range of applications, including computer engineering and robotics. M. C. Thar et al. [1] proposed a pose assessment method for yoga using the technique of pose detection developed by Thar. It works on the principle of self-examining yoga practise techniques with the use of the pose detection method. This paper basically presents the idea of a performance evaluation system as a yoga pose tutor system. So it offers a mechanism to locate yoga poses and to utilise pose discovery so that it can assist in the self-mastery of yoga.

Santosh Kumar Yadav et al. [2] proposed real-time yoga pose recognition using deep learning. This paper suggests a high-intensity hybrid mastering version for revealing yoga poses in real-time videos. This paper primarily proposes an assistant yoga app based on human key acquisition techniques for video chat. This in-depth hybrid mastery model has provided a concept to employ CNN and LSTM to disclose yoga in real-time movies in which to extract capabilities from the critical state of anyone determined in open stance and viewed using LSTM to supply transitory forecasts. A CNN network primarily processes images, whereas an LSTM is a model that is commonly used for time series predictions.

A deep learning model for yoga pose monitoring was proposed by Debabrata Swain et al. [3]. It suggested improving the work that had already been done and comparing it to other projects with a similar framework. The use of additional crucial points in the suggested work produces better and more dependable vaticination quality. Additionally, the batch normalisation (BN) subcaste is applied after the powerhouse subcaste for quicker confluence; this preface of the BN subcaste makes sure that the input pixels have an analogous distribution and speeds up the training. Additionally, the Adadelta optimizer is slowed down and constrained by accretive confluence. M. U. Islam et al. [4] proposed a method for recognising yoga postures by revealing human joint points in real time using a technique called Microsoft Kinect. The proposed system can track the precision of various yoga postures as well as the movement of human body parts, making it easier for the user to practise yoga. Microsoft Kinect will recognise various joint points on the human body in real time and compute various angles from those joint points to determine how accurately a user is performing a specific yoga posture.

Rutuja Gajbhiye et al. [5] proposed an artificial intelligence-based estimation of human pose as well as pose correction and estimation. By identifying key points on the user's body, this model generates a virtual "skeleton" in 2D or 3D dimensions. The virtual skeleton is examined for flaws in the exercise using geometry-based rules or other techniques. The user is given an explanation of the errors made as well as suggestions for correcting them. Machine learning techniques for the detection of different yoga asanas were proposed by Rutuja Jagtap et al. [6]. To correctly identify the yoga stance and make predictions in accordance with it, the system worked on a total of eight asans, including Vrikshasan, Bhadrasan, Gomukhasan, Shavasan, Chakrasan, Shrishasan, Sarvangasana, and Dhanurasan.

III. FUNCTIONALITIES AND FEATURES

A. *Functionalities*

- First, the purpose of this model is to accurately identify the yoga poses performed by trainees using a machine vision system to create uniform and useful yoga positions.
- Second, the goal of our research model is to integrate all of the features from various platforms related to yoga pose detection applications on a single platform. This will be a more effective and user-interactive system for users, providing them with more options on a single platform rather than multiple ones.
- Third, we are trying to combine the benefits of both physical trainers and AI tutors on one platform to provide users with more options based on their needs.
- Moreover, this study intends to be as user-interactive as possible. When we say “user-interactive,” we mean that the user should feel at ease using our platform and receive written instructions as they move through it so that they can easily use the app and its corresponding features. Furthermore, the user will receive all information relevant to the specific yoga pose that he or she is performing, including the name of the pose, its benefits, and the posture's actual correct execution.
- By directly integrating with calendar reminders, this model enables users to conveniently schedule their diet plans.
- It will also automatically count the number of times the user performs push-ups, uses dumbbells, and many other yoga poses that require counting the number of times the user performs that task. There will also be an option for users to set alarms when they reach a certain limit as part of their dieting plans, allowing them to exercise with peace of mind.

B. *Features*

1) Intensive yoga instruction: This model is very important because it gives practitioners instant feedback and lets them fix their alignment. As a result, practitioners can practise yoga safely and effectively [7].

2) Individualized Feedback: Practitioners will receive individualised feedback based on this model, allowing them to concentrate on their strengths and weaknesses [8].

3) Consistency: The practitioner can use this model to make sure they practise yoga poses on a regular basis, gaining a thorough understanding of each one [9].

4) Preventing Injuries: This model will educate the practitioner about the dangers of incorrect postures and the injuries they can cause. Therefore, improper practise-related injuries can be avoided with the assistance of this model [10].

IV. METHODOLOGY

A. *OpenCV and MediaPipe architecture*

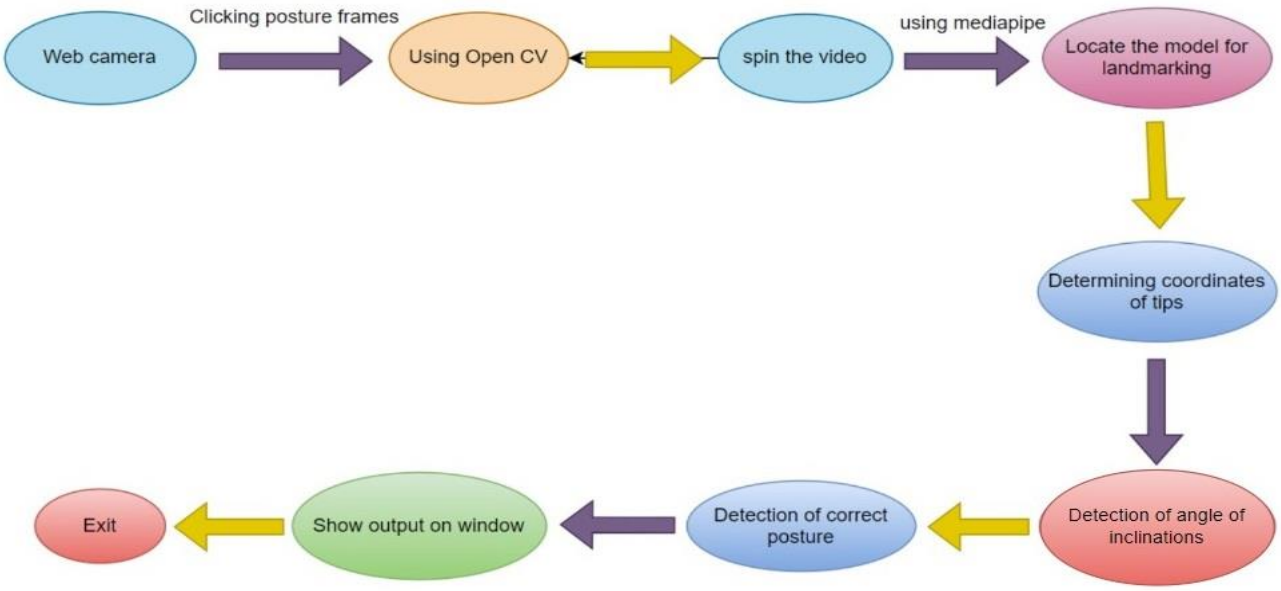

Fig. 2 Open CV and MediaPipe Architecture

1) MediaPipe: MediaPipe is an open-source framework created by Google in June 2019 in order to provide machine learning solutions for developing models and applications that use webcams or audio to detect objects in real time [11]. MediaPipe provides access to a variety of prominent machine learning frameworks that were created with portable hardware restrictions in mind. These versions include hand tracing, pose tracking, facial mesh tracking, object detection in 2D and 3D, and pose estimation.

2) OpenCV: OpenCV is an open-source computer vision library written in C++, but it also provides APIs for other programming languages such as Python and Java. OpenCV is used to process real-time images, videos, and even handwritten manuscripts [12]. This library helps in the detection of objects such as faces, images, and videos, and it also includes other Python libraries such as numpy and matplotlib that provide additional features for working with applications.

Here, Fig. 2 depicts the integrated operation of the OpenCV and MediaPipe architectures. The model operates as follows:

- In this model, all relevant libraries and dependencies are loaded first.
- The next step is data preprocessing, in which the person is detected if it is within the detection frame of the webcam.
- Then, based on the pose landmarks, such as joints, the yoga poses are identified. Basically, this model uses 33 landmarks to distinguish between various joints, which helps in recognising poses.
- We can also calculate the angle of a specific landmark to aid in predicting the correct posture, and our model will provide the correct angle as feedback to the user so that the correct posture can be practised.
- There is also a live counter that enhances user interaction by providing a live countdown for the gym activity he is doing; for example, if a user is doing pushups, it will give the number of pushups the user has done while also detecting the user's posture and providing live feedback.

B. Proposed Model

Yoga pose detection is an AI-based yoga tracker model that is based on computer vision. Its primary aim is to classify yoga poses from images and videos of people or to suggest some feedback about their yoga postures to those who are performing live in front of the device on which this model is installed. The basic algorithm used for this model is based on human posture estimation, which means that when practitioners perform any pose in front of this model, the camera captures their movements, splits the angular coordinates into individual frames, and forms a virtual “skeleton” in 2D or 3D dimensions. This virtual skeleton is used to detect the body joints and angles of a person in order to identify errors in the exercise by comparing the angles with the preinstalled dataset present in it. This feature is added by using the MediaPipe pose estimation technique. Moreover, while creating this model, we used some classification algorithms like MediaPipe, which will help in classifying the yoga poses based on angles of inclination and offset distance.

This system first captures input using the webcam and Photoshops the framework from a yoga pose that was captured during live video chatting. Then it is transformed into an RGB frame (red, green, and blue frame) so that it can be seen on a computer's digital display screen and be used to capture multiple-channel images. The system will then perform some internal preprocessing. Then, it will start locating the key points in the posture and simultaneously calculating the offset distance. The system will alert you if these strategically placed key points are not properly coordinated; otherwise, it will only reveal the offset distance. It will now begin determining the body posture's inclination angle. The body posture will then be compared to the ideal one using the classification algorithm, or MediaPipe, in this case. If it is a good posture, a success message with positive feedback will be shown; if it is not, it will give negative feedback and restart. The overall framework of our suggested model, or how it internally functions to detect the right poses and provide feedback accordingly, is shown in Fig. 3.

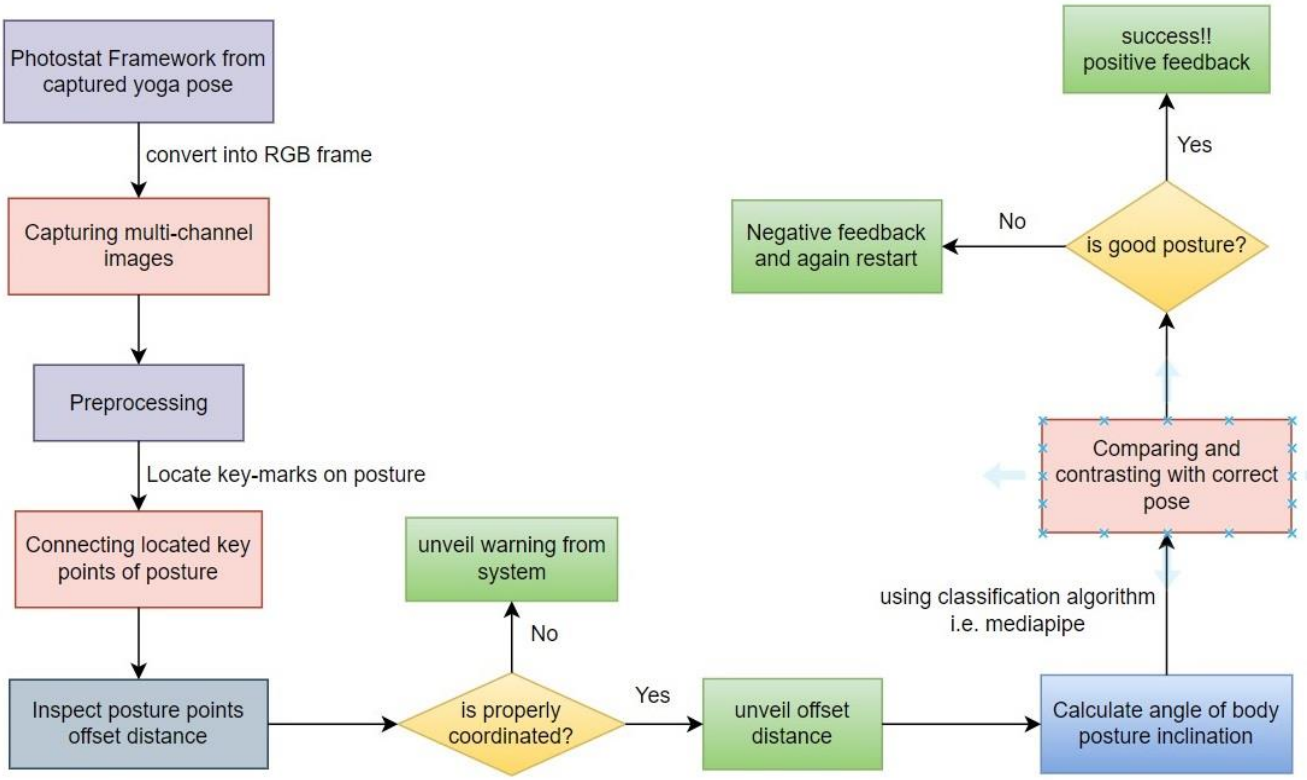

Fig. 3. Framework of proposed model

We must have 4 GB of RAM and a 64-bit operating system to run our system, which is implemented in python using the OpenCV library. This system can be used as an app or integrated with a platform for tracking yoga. In a nutshell, this system relies on deep learning classifiers and machine learning classifiers for particular tasks. As a result, this model is a useful resource for both yoga students and teachers. It enables instructors to give more individualised feedback and guidance while also enabling practitioners to monitor their progress and make sure they are performing poses properly.

V. RESULTS AND DISCUSSIONS

The potential of a model that identifies yoga postures is very promising for the future. As yoga gains popularity and wearable technology becomes more prevalent, there will be an increasing demand for precise and effective yoga pose detection models [13]. Personalised yoga pose recommendations could be included in the model. This would enable a yoga pose detection model to observe and evaluate a person's yoga practise and suggest customised sequences of poses based on their level of expertise and physical condition. Another feature is that users could connect directly with a yoga instructor while performing yoga in front of the yoga pose detection model, so instructors could correct their students' poses in real-time. Additionally, the model could help prevent injuries and encourage safe and healthy yoga practises. A large dataset of yoga poses and movements could be collected and analysed to better understand the biomechanics and health benefits of yoga. Furthermore, a pose detection model could be integrated into fitness equipment such as yoga mats or mirrors, allowing users to receive real-time feedback on their poses and movements. Ultimately, a yoga pose detection model has the potential to significantly enhance the practise of yoga and promote better health and wellbeing for individuals [14].

To detect yoga poses, MediaPipe is one of the best frameworks that works more efficiently than any other technique using a single webcam [15]. First, the object or person is detected, and then the model estimates the curls by using landmarks that assist the model in detecting the joints. At the same time, the pink colour line is visible, which is used to differentiate the object from the background. There is also a blue counter box that counts the pushups done while doing gym activities.

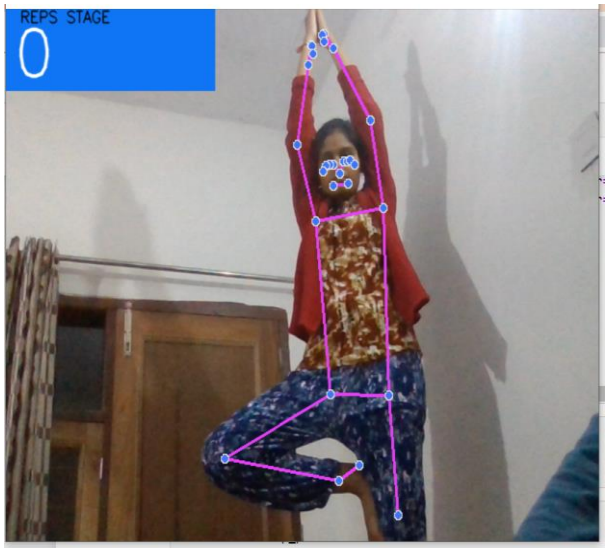

Fig. 4. Yoga Pose 1(Vraksasana)

As shown in Fig. 4, when the user is performing Vraksasana, all the joints are displayed in pink, and the person’s pose is recognised.

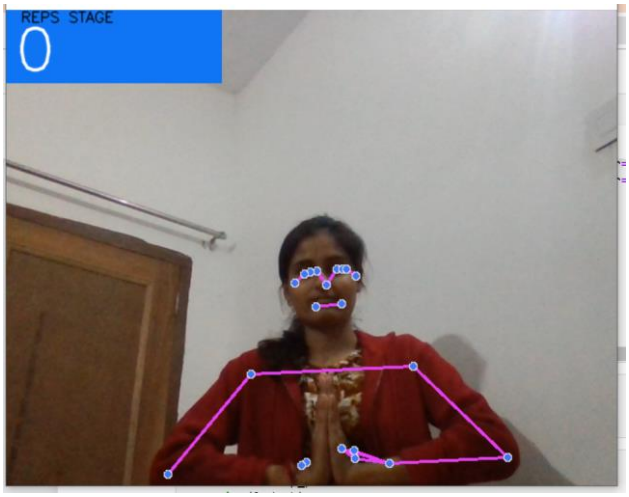

Fig. 5. Yoga Pose 2 (Padmasana)

As seen in Fig. 5, the person's posture is recognised, and the curls are depicted using the pink colour to distinguish them from the frame.

VI. CONCLUSION

The entire methodology used in this study is based on deep learning, which is used to both identify incorrect yoga postures and direct the user if something is incorrect. We encountered many difficulties while developing this model, such as identifying the correct yoga poses due to variations in conditions, clothing, and body shape, as well as the complexity of some postures. Despite these obstacles, new research on the subject has revealed promising accuracy rates in detecting and recognising yoga poses. The technique produces results based on the angles between joints in an AI-created virtual skeleton. If any key points are rotated slightly, the angles change, and the feedback generation system produces feedback. With further advances in computer vision and machine learning, we can expect to see more progress in yoga pose detection and its applications in areas such as yoga training, fitness monitoring, and healthcare.

More research is required to investigate the potential of these systems for various types of yoga and practitioners of varying levels. We can have a recommendation system for different yoga exercises that will recommend the types of yoga we should do at the time, as any yoga asana practised at the wrong time will not give good results and can have some adverse effects on our body. There can be an AI assistant in the project that will help the user by giving voice directions and feedback for different yoga postures practised. In order to enhance the lively interaction with the user, we can add microphone functionality that will give live countdowns using voice recognition to track the gym activities of a user.

REFERENCES

[1] M. C. Thar, K. Z. N. Winn, and N. Funabiki, “A Proposal of Yoga Pose Assessment Method Using Pose Detection for Self-Learning,” 2019 International Conference on Advanced Information Technologies (ICAIT), Yangon, Myanmar, 2019, pp. 137-142.

[2] Santosh Kumar Yadav, Amitojdeep Singh, Abhishek Gupta, and Jagdish Lal Raheja, “Real-time Yoga recognition using deep learning,” Neural Computing and Applications, vol. 31, pp. 9349–9361, 2019.

[3] Debabrata Swain, Santosh Satapathy, Biswaranjan Acharya, Madhu Shukla, Vassilis C. Gerogiannis, Andreas Kanavos, and Dimitris Giakovis, “Deep Learning Models for Yoga Pose Monitoring,” Algorithms, vol. 15, Article No. 403, 2022.

[4] M. U. Islam, H. Mahmud, F. B. Ashraf, I. Hossain, and M. K. Hasan, “Yoga posture recognition by detecting human joint points in real time using microsoft kinect,” 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, Bangladesh, 2017, pp. 668-673.

[5] Rutuja Gajbhiye, Snehal Jarag, Pooja Gaikwad, and Shweta Koparde, “AI Human Pose Estimation: Yoga Pose Detection and Correction,” International Journal of Innovative Science and Research Technology, vol.7, pp. 1649-1658, 2022.

[6] Rutuja Jagtap, Monali Zanzane, and Rutuja Patil, “Yoga Pose Detection Using Machine Learning,” International Research Journal of Modernization in Engineering Technology and Science, vol. 4, pp. 1753-1756, 2022.

[7] F. Rishan, B. De Silva, S. Alawathugoda, S. Nijabdeen, L. Rupasinghe, and C. Liyanapathirana, “Infinity Yoga Tutor: Yoga Posture Detection and Correction System,” 2020 5th International Conference on Information Technology Research (ICITR), Moratuwa, Sri Lanka, 2020, pp. 1-6.

[8] J. Palanimeera and K. Ponmozhi, “Classification of yoga pose using machine learning techniques,” Materials Today: Proceedings, vol. 37, pp. 2930-2933, 2021

[9] Y. Agrawal, Y. Shah, and A. Sharma, “Implementation of Machine Learning Technique for Identification of Yoga Poses,” 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT), Gwalior, India, 2020, pp. 40-43.

[10] O. Tarek, O. Magdy and A. Atia, "Yoga Trainer for Beginners Via Machine Learning," 2021 9th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC), Alexandria, Egypt, 2021, pp. 75-78.

[11] S. Liaqat, K. Dashtipour, K. Arshad, K. Assaleh, and N. Ramzan, “A Hybrid Posture Detection Framework: Integrating Machine Learning and Deep Neural Networks,” IEEE Sensors Journal, vol. 21, pp. 9515-9522, 2021.

[12] M.V.D. Prasad, K. Durga Bhavani, S. M. Tharun Kumar, P.V.V. Kishore, M. Teja Kiran Kumar, E. Kiran Kumar, and D. Anil Kumar, “Machine learning based 2D pose estimation model for human action recognition using geometrical maps” International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 8, pp. 1127-1130, 2019.

[13] Chirumamilla Nagalakshmi and Snehasis Mukherjee, “Classification of Yoga Asanas from a Single Image by Learning the 3D View of Human Poses,” in Digital Techniques for Heritage Presentation and Preservation, Mukhopadhyay, J., Sreedevi, I., Chanda, B., Chaudhury, S., Namboodiri, V.P. Eds., Springer, Cham, 2021, pp. 37-49.

[14] Chhaihuoy Long, Eunhye Jo, and Yunyoung Nam, “Development of a yoga posture coaching system using an interactive display based on transfer learning,” The Journal of Supercomputing, vol. 78, pp. 5269–5284, 2022.

[15] Utkarsh Bahukhandi and Dr. Shikha Gupta,“Yoga Pose Detection And Classification Using Machine Learning Techniques” International Research Journal of Modernization in Engineering Technology and Science, vol. 3, pp. 186-191, 2021.